

Weekly Report

Yuxin Ma

10.08.2012 - 10.14.2012

Intro

Before Wednesday I was preparing for the seminar presentation. This week the active learning-based visual analysis project and the heterogeneous network visualization project are still going on. It seems that some bottlenecks have been met during these days.

Research

- **Active Learning** Actually I have confused about the application scenario for couples of weeks. This week I revised two surveys [1] and [2] for more application scenarios in order to try something bottom-up. At present the “noisy oracle” problem seems to be one of the directions that visual analysis approaches can be applied. There are some points related to this topic:
 - Human as oracles can sometimes mis-tag some queries generated from model and lead to uncertain process.
 - This alternative tagging results construct different branches during the whole iterative learning process.
 - Analysts can perform some low-cost experiments of different tagging strategies to see different “growing” results.
 - The whole process can be considered as a themeriver describing the whole iterative learning process.

I am going to search some active clustering algorithms as an application for a design next week.

- **Rough idea on Taobao categorical data** This week I got the Taobao categorical data from Xinxin Huang. There are some rough ideas about analysing large categorical data:
 - Latent Diriclet Allocation can be applied to extract hidden topics of massive tags, however no semantic information is available from this approach. After LDA applied, we can analyse these “tag topics” to extract meanings behind them.
 - Interactive modification of categories is available(just like what Xinxin Huang is doing now).
 - The temporal changes of tags can be explored via some flowmap-based approaches; semantic meanings of tags might be alternative in different time periods.

Practice & Skills

- **openFramework's Migration to Qt** This week I did an exploration on migrating openFrameworks into Qt in order to use native UI components in openFrameworks interface. This work is still ongoing.

Miscellaneous

None.

Plan for Next Week

- For active learning project, some applications will be investigated to support the “noisy oracle” scenario.

References

- [1] B. Settles, *Active Learning (Synthesis Lectures on Artificial Intelligence and Machine Learning)*. Morgan & Claypool Publishers, July 2012.
- [2] B. Settles, “Active learning literature survey,” *University of Wisconsin, Madison*, 2010.